



IFWO

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/726,645

TIME: 08:36:29

Input Set : N:\Crf3\RULE60\10726645.raw.txt

Output Set: N:\CRF4\09282004\J726645.raw

```

1 <110> APPLICANT: Hastings, et al.
2 <120> TITLE OF INVENTION: HUMAN OSTEOCLAST-DERIVED CATHEPSIN
3 <130> FILE REFERENCE: PF107D5
4 <140> CURRENT APPLICATION NUMBER: US/10/726,645
5 <141> CURRENT FILING DATE: 2003-12-04
6 <150> PRIOR APPLICATION NUMBER: US/10/114,464
7 <151> PRIOR FILING DATE: 2002-04-03
8 <150> PRIOR APPLICATION NUMBER: 08/553,125
9 <151> PRIOR FILING DATE: 1995-11-07
10 <150> PRIOR APPLICATION NUMBER: 08/208,007
11 <151> PRIOR FILING DATE: 1994-03-08
12 <160> NUMBER OF SEQ ID NOS: 14
13 <170> SOFTWARE: PatentIn Ver. 2.1
15 <210> SEQ ID NO: 1
16 <211> LENGTH: 1619
17 <212> TYPE: DNA
18 <213> ORGANISM: Homo sapiens
19 <220> FEATURE:
20 <221> NAME/KEY: CDS
21 <222> LOCATION: (21)..(1009)
22 <220> FEATURE:
23 <221> NAME/KEY: mat_peptide
24 <222> LOCATION: (365)..(1009)
25 <220> FEATURE:
26 <221> NAME/KEY: sig_peptide
27 <222> LOCATION: (21)..(1009)
28 <400> SEQUENCE: 1
29      tcagatttcc atcagcagg atg tgg ggg ctc aag gtt ctg ctg cta cct gtg   52
30      Met Trp Gly Leu Lys Val Leu Leu Leu Pro Val
> 31      -110                                -105
32      gtg agc ttt gct ctg tac cct gag gag ata ctg gac acc cac tgg gag   100
33      Val Ser Phe Ala Leu Tyr Pro Glu Glu Ile Leu Asp Thr His Trp Glu
34      -100                                -95                                -90
35      cta tgg aag aag acc cac agg aag caa tat aac aac aag gtg gat gaa   148
36      Leu Trp Lys Lys Thr His Arg Lys Gln Tyr Asn Asn Lys Val Asp Glu
37      -85                                -80                                -75
38      atc tct cgg cgt tta att tgg gaa aaa aac ctg aag tat att tcc atc   196
39      Ile Ser Arg Arg Leu Ile Trp Glu Lys Asn Leu Lys Tyr Ile Ser Ile
40      -70                                -65                                -60
41      cat aac ctt gag gct tct ctt ggt gtc cat aca tat gaa ctg gct atg   244
42      His Asn Leu Glu Ala Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met
43      -55                                -50                                -45
44      aac cac ctg ggg gac atg acc agt gaa gag gtg gtt cag aag atg act   292

```



## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/726,645

TIME: 08:36:29

Input Set : N:\Crif3\RULE60\10726645.raw.txt

Output Set: N:\CRF4\09282004\J726645.raw

```

45      Asn His Leu Gly Asp Met Thr Ser Glu Glu Val Val Gln Lys Met Thr
46      -40                      -35                      -30                      -25
47      gga ctc aaa gta ccc ctg tct cat tcc cgc agt aat gac acc ctt tat   340
48      Gly Leu Lys Val Pro Leu Ser His Ser Arg Ser Asn Asp Thr Leu Tyr
49                      -20                      -15                      -10
50      atc cca gaa tgg gaa ggt aga gcc cca gac tct gtc gac tat cga aag   388
51      Ile Pro Glu Trp Glu Gly Arg Ala Pro Asp Ser Val Asp Tyr Arg Lys
52                      -5                      -1 1                      5
53      aaa gga tat gtt act cct gtc aaa aat cag ggt cag tgt ggt tcc tgt   436
54      Lys Gly Tyr Val Thr Pro Val Lys Asn Gln Gly Gln Cys Gly Ser Cys
55      10                      15                      20
56      tgg gct ttt agc tct gtg ggt gcc ctg gag ggc caa ctc aag aag aaa   484
57      Trp Ala Phe Ser Ser Val Gly Ala Leu Glu Gly Gln Leu Lys Lys Lys
58      25                      30                      35                      40
59      act ggc aaa ctc tta aat ctg agt ccc cag aac cta gtg gat tgt gtg   532
60      Thr Gly Lys Leu Leu Asn Leu Ser Pro Gln Asn Leu Val Asp Cys Val
61                      45                      50                      55
62      tct gag aat gat ggc tgt gga ggg ggc tac atg acc aat gcc ttc caa   580
63      Ser Glu Asn Asp Gly Cys Gly Gly Tyr Met Thr Asn Ala Phe Gln
64                      60                      65                      70
65      tat gtg cag aag aac cgg ggt att gac tct gaa gat gcc tac cca tat   628
66      Tyr Val Gln Lys Asn Arg Gly Ile Asp Ser Glu Asp Ala Tyr Pro Tyr
67      75                      80                      85
68      gtg gga cag gaa gag agt tgt atg tac aac cca aca ggc aag gca gct   676
69      Val Gly Gln Glu Glu Ser Cys Met Tyr Asn Pro Thr Gly Lys Ala Ala
70      90                      95                      100
71      aaa tgc aga ggg tac aga gag atc ccc gag ggg aat gag aaa gcc ctg   724
72      Lys Cys Arg Gly Tyr Arg Glu Ile Pro Glu Gly Asn Glu Lys Ala Leu
73      105                      110                      115                      120
74      aag agg gca gtg gcc cga gtg gga cct gtc tct gtg gcc att gat gca   772
75      Lys Arg Ala Val Ala Arg Val Gly Pro Val Ser Val Ala Ile Asp Ala
76      125                      130                      135
77      agc ctg acc tcc ttc cag ttt tac agc aaa ggt gtg tat tat gat gaa   820
78      Ser Leu Thr Ser Phe Gln Phe Tyr Ser Lys Gly Val Tyr Tyr Asp Glu
79      140                      145                      150
80      agc tgc aat agc gat aat ctg aac cat gcg gtt ttg gca gtg gga tat   868
81      Ser Cys Asn Ser Asp Asn Leu Asn His Ala Val Leu Ala Val Gly Tyr
82      155                      160                      165
83      gga atc cag aag gga aac aag cac tgg ata att aaa aac agc tgg gga   916
84      Gly Ile Gln Lys Gly Asn Lys His Trp Ile Ile Lys Asn Ser Trp Gly
85      170                      175                      180
86      gaa aac tgg gga aac aaa gga tat atc ctc atg gct cga aat aag aac   964
87      Glu Asn Trp Gly Asn Lys Gly Tyr Ile Leu Met Ala Arg Asn Lys Asn
88      185                      190                      195                      200
89      aac gcc tgt ggc att gcc aac ctg gcc agc ttc ccc aag atg tga   1009
90      Asn Ala Cys Gly Ile Ala Asn Leu Ala Ser Phe Pro Lys Met
91      205                      210
92      ctccagccag ccaaaccat cctgctcttc catttcttcc acgatgggtgc agtghtaacga 1069
93      tgcacttttg aaggaggattg gtgtgctatt tttgaagcag atgtggtgat actgagattg 1129

```

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/726,645

TIME: 08:36:29

Input Set : N:\Crif3\RULE60\10726645.raw.txt

Output Set : N:\CRF4\09282004\J726645.raw

```

94      tctgttcagt ttccccattt gtttgtgctt caaatgatec ttcctacttt gcttctctcc 1189
95      acccatgacc tttttcactg tggccatcag gactttcccc tgacagctgt gtactcttag 1249
96      gctaagagat gtgactacag cctgcccctg actgtgttgt cccagggctg atgctgtaca 1309
97      ggtacaggct ggagattttc acataggta gattctcatt cacgggacta gttagcttta 1369
98      agcaccctag aggactaggg taatctgact tctcacttcc taagtccct tctatattcct 1429
99      caaggtagaa atgtctatgt tttctactcc aattcataaa tctattcata agtctttggt 1489
100     acaagtttac atgataaaaa gaaatgtgat ttgtcttccc ttctttgcac ttttgaaata 1549
101     aagtatttat ctctgtctta cagttaata aatagcatct agtacacatt caaaaaaaaaa 1609
102     aaaaaaaaaa 1619
104 <210> SEQ ID NO: 2
105 <211> LENGTH: 329
106 <212> TYPE: PRT
107 <213> ORGANISM: Homo sapiens
108 <400> SEQUENCE: 2
109     Met Trp Gly Leu Lys Val Leu Leu Leu Pro Val Val Ser Phe Ala Leu
110     -115 -110 -105 -100
111     Tyr Pro Glu Glu Ile Leu Asp Thr His Trp Glu Leu Trp Lys Lys Thr
112     -95 -90 -85
113     His Arg Lys Gln Tyr Asn Asn Lys Val Asp Glu Ile Ser Arg Arg Leu
114     -80 -75 -70
115     Ile Trp Glu Lys Asn Leu Lys Tyr Ile Ser Ile His Asn Leu Glu Ala
116     -65 -60 -55
117     Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met Asn His Leu Gly Asp
118     -50 -45 -40
119     Met Thr Ser Glu Glu Val Val Gln Lys Met Thr Gly Leu Lys Val Pro
120     -35 -30 -25 -20
121     Leu Ser His Ser Arg Ser Asn Asp Thr Leu Tyr Ile Pro Glu Trp Glu
122     -15 -10 -5
123     Gly Arg Ala Pro Asp Ser Val Asp Tyr Arg Lys Lys Gly Tyr Val Thr
124     -1 1 5 10
125     Pro Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ser
126     15 20 25
127     Val Gly Ala Leu Glu Gly Gln Leu Lys Lys Lys Thr Gly Lys Leu Leu
128     30 35 40 45
129     Asn Leu Ser Pro Gln Asn Leu Val Asp Cys Val Ser Glu Asn Asp Gly
130     50 55 60
131     Cys Gly Gly Gly Tyr Met Thr Asn Ala Phe Gln Tyr Val Gln Lys Asn
132     65 70 75
133     Arg Gly Ile Asp Ser Glu Asp Ala Tyr Pro Tyr Val Gly Gln Glu Glu
134     80 85 90
135     Ser Cys Met Tyr Asn Pro Thr Gly Lys Ala Ala Lys Cys Arg Gly Tyr
136     95 100 105
137     Arg Glu Ile Pro Glu Gly Asn Glu Lys Ala Leu Lys Arg Ala Val Ala
138     110 115 120 125
139     Arg Val Gly Pro Val Ser Val Ala Ile Asp Ala Ser Leu Thr Ser Phe
140     130 135 140
141     Gln Phe Tyr Ser Lys Gly Val Tyr Tyr Asp Glu Ser Cys Asn Ser Asp
142     145 150 155
143     Asn Leu Asn His Ala Val Leu Ala Val Gly Tyr Gly Ile Gln Lys Gly

```

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/726,645

TIME: 08:36:29

Input Set : N:\Crif3\RULE60\10726645.raw.txt

Output Set: N:\CRF4\09282004\J726645.raw

```

144          160          165          170
145      Asn Lys His Trp Ile Ile Lys Asn Ser Trp Gly Glu Asn Trp Gly Asn
146          175          180          185
147      Lys Gly Tyr Ile Leu Met Ala Arg Asn Lys Asn Asn Ala Cys Gly Ile
148          190          195          200          205
149      Ala Asn Leu Ala Ser Phe Pro Lys Met
150          210
152 <210> SEQ ID NO: 3
153 <211> LENGTH: 26
154 <212> TYPE: DNA
155 <213> ORGANISM: Homo sapiens
156 <400> SEQUENCE: 3
157      gctaaggatc ctgggggctc aaggtt
159 <210> SEQ ID NO: 4
160 <211> LENGTH: 26
161 <212> TYPE: DNA
162 <213> ORGANISM: Homo sapiens
163 <400> SEQUENCE: 4
164      gctaattctag atcacatctt ggggaa
166 <210> SEQ ID NO: 5
167 <211> LENGTH: 14
168 <212> TYPE: PRT
169 <213> ORGANISM: Homo sapiens
170 <400> SEQUENCE: 5
171      Ala Ile Asp Ala Ser Leu Thr Ser Phe Gln Phe Tyr Ser Lys
172          1          5          10
174 <210> SEQ ID NO: 6
175 <211> LENGTH: 11
176 <212> TYPE: PRT
177 <213> ORGANISM: Homo sapiens
178 <400> SEQUENCE: 6
179      Tyr Asp Glu Ser Cys Asn Ser Asp Asn Leu Asn
180          1          5          10
182 <210> SEQ ID NO: 7
183 <211> LENGTH: 329
184 <212> TYPE: PRT
185 <213> ORGANISM: Homo sapiens
186 <400> SEQUENCE: 7
187      Met Trp Gly Leu Lys Val Leu Leu Leu Pro Val Val Ser Phe Ala Leu
188          1          5          10          15
189      His Pro Glu Glu Ile Leu Asp Thr Gln Trp Glu Leu Trp Lys Lys Thr
190          20          25          30
191      Tyr Ser Lys Gln Tyr Asn Ser Lys Val Asp Glu Ile Ser Arg Arg Leu
192          35          40          45
193      Ile Trp Glu Lys Asn Leu Lys His Ile Ser Ile His Asn Leu Glu Ala
194          50          55          60
195      Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met Asn His Leu Gly Asp
196          65          70          75          80
197      Met Thr Ser Glu Glu Val Val Gln Lys Met Thr Gly Leu Lys Val Pro

```

## RAW SEQUENCE LISTING

DATE: 09/28/2004

PATENT APPLICATION: US/10/726,645

TIME: 08:36:29

Input Set : N:\Crf3\RULE60\10726645.raw.txt

Output Set: N:\CRF4\09282004\J726645.raw

```

198          85          90          95
199  Pro Ser Arg Ser His Ser Asn Asp Thr Leu Tyr Ile Pro Asp Trp Glu
200          100          105          110
201  Gly Arg Thr Pro Asp Ser Ile Asp Tyr Arg Lys Lys Gly Tyr Val Thr
202          115          120          125
203  Pro Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ser
204          130          135          140
205  Val Gly Ala Leu Glu Gly Gln Leu Lys Lys Lys Thr Gly Lys Leu Leu
206          145          150          155          160
207  Asn Leu Ser Pro Gln Asn Leu Val Asp Cys Val Ser Glu Asn Tyr Gly
208          165          170          175
209  Cys Gly Gly Gly Tyr Met Thr Asn Ala Phe Gln Tyr Val Gln Arg Asn
210          180          185          190
211  Arg Gly Ile Asp Ser Glu Asp Ala Tyr Pro Tyr Val Gly Gln Asp Glu
212          195          200          205
213  Ser Cys Met Tyr Asn Pro Thr Gly Lys Ala Ala Lys Cys Arg Gly Tyr
214          210          215          220
215  Arg Glu Ile Pro Glu Gly Asn Glu Lys Ala Leu Lys Arg Ala Val Ala
216          225          230          235          240
217  Arg Val Gly Pro Val Ser Val Ala Ile Asp Ala Ser Leu Thr Ser Phe
218          245          250          255
219  Gln Phe Tyr Ser Lys Gly Val Tyr Tyr Asp Glu Asn Cys Ser Ser Asp
220          260          265          270
221  Asn Val Asn His Ala Val Leu Ala Val Gly Tyr Gly Ile Gln Lys Gly
222          275          280          285
223  Asn Lys His Trp Ile Ile Lys Asn Ser Trp Gly Glu Ser Trp Gly Asn
224          290          295          300
225  Lys Gly Tyr Ile Leu Met Ala Arg Asn Lys Asn Asn Ala Cys Gly Ile
226          305          310          315          320
227  Ala Asn Leu Ala Ser Phe Pro Lys Met
228          325
230 <210> SEQ ID NO: 8
231 <211> LENGTH: 331
232 <212> TYPE: PRT
233 <213> ORGANISM: Homo sapiens
234 <400> SEQUENCE: 8
235  Met Lys Arg Leu Val Cys Val Leu Leu Val Cys Ser Ser Ala Val Ala
236      1          5          10          15
237  Gln Leu His Lys Asp Pro Thr Leu Asp His His Trp His Leu Trp Lys
238          20          25          30
239  Lys Thr Tyr Gly Lys Gln Tyr Lys Glu Lys Asn Glu Glu Ala Val Arg
240          35          40          45
241  Arg Leu Ile Trp Glu Lys Asn Leu Lys Phe Val Met Leu His Asn Leu
242          50          55          60
243  Glu His Ser Met Gly Met His Ser Tyr Asp Leu Gly Met Asn His Leu
244          65          70          75          80
245  Gly Asp Met Thr Ser Glu Glu Val Met Ser Leu Thr Ser Ser Leu Arg
246          85          90          95
247  Val Pro Ser Gln Trp Gln Arg Asn Ile Thr Tyr Lys Ser Asn Pro Asn

```

VERIFICATION SUMMARY

DATE: 09/28/2004

PATENT APPLICATION: US/10/726,645

TIME: 08:36:30

Input Set : N:\Crf3\RULE60\10726645.raw.txt

Output Set: N:\CRF4\09282004\J726645.raw

31 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1